**AP Stat- Unit 2 In-class review problem - ANSWERS**

Below is data on the years of education versus the years spent in jail by a sample of 20 – 40 year old men. *For this question, round all numbers to3 decimal places.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Education (Yrs)** | **Jail Time (Yrs)** |  | **Education (Yrs)** | **Jail Time (Yrs)** |
| 24 | 0 |  | 10 | 5.2 |
| 20 | 2.1 |  | 28 | 0.1 |
| 12 | 5.2 |  | 5 | 8.7 |
| 13 | 3.6 |  | 8 | 8.9 |
| 20 | 0.5 |  | 9 | 7.6 |
| 21 | 1 |  | 12 | 2.3 |
| 10 | 2.2 |  | 14 | 4.5 |
| 6 | 6.5 |  | 15 | 2.1 |
| 8 | 7 |  | 17 | 1.3 |
| 10 | 4 |  | 21 | 0.4 |
| 16 | 2.5 |  | 23 | 0.9 |
| 18 | 1.6 |  | 7 | 9.1 |

1. Determine the explanatory and response variables

Education = explanatory Jail time = response

1. Sketch a scatterplot of the data. Describe the scatterplot.



The plot is linear, negative, and moderately strong. No outliers.

1. Find the equation of the LSR line and the correlation coefficient. Sketch the LSR line on your scatterplot from (a).

= 9.6 – 0.412(X) or = 9.6 – 0.412(education) r = --0.8775

1. Use the model to predict the number of years in jail for someone with 18 years of education.

= 9.6 – 0.412(18)

Jail time = 2.184 years

1. Calculate the residual for the prediction in part (d)

According to the original data given, someone with 18years of education spent 1.6 years in jail.

Residual = actual – predicted = 1.6 – 2.184 = -0.584 years

1. Is this prediction an overestimate or an underestimate? Why?

Overestimate. The line overestimated the amount of time that a person with 18 years of education would spend in jail. The line predicted 2.184 years when it was actually 1.6 years.

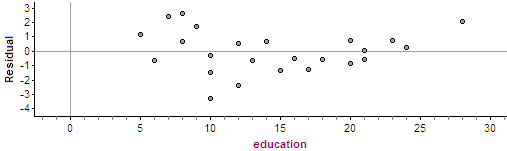
1. Interpret the slope of the LSR line in a complete sentence.

For every year increase in education, on average a person spends 0.412 years LESS in jail.

1. **Given that** a person has spent 5 years in jail, how many years of education would you predict they have had? (SHOW WORK)

5 = 9.6 – 0.412(education)

Education = 11.165 years of education

1. Sketch the residual plot.   
   
2. What does the residual plot in part (i) tell us about our linear model? Justify.

Since there seems to be a pattern in the residual plot, we can conclude that our linear model is not the best fit model for our data. There is another model that would fit our data better.

***\*\*BE CAREFUL: DO NOT SAY THAT THE LINEAR MODEL IS A BAD FIT FOR THE DATA!!!***

1. Find the coefficient of determination (r2) and interpret it.

r2 = 0.77

77% of the change in jail time is explained by the change in education. OR

77% of the change in jail time is explained by the LSR line on education

1. List some confounding/lurking variables in this situation.

Where the person lives, their family history, etc.